Instructional Course:
Videolaryngo-Stroboscopy and Phonetography –
Basic Tools for Diagnostics and Documentation in the Voice Clinic

JÜRGEN WENDLER, TADEUS NAWKA, DIRK VERGES

1. Videolaryngo-Stroboscopy

1.1 Introduction

1.1.1 The Stroboscopic Principle

The stroboscopic effect is based on an optical illusion arising from the persistence of visual impressions. According to the Talbot law, a sequence of individual frames, presented at a critical flicker frequency, appears as a continuously moving picture. Rapid, periodic consecutive motions (as for example vibrational movements) that cannot be resolved by the human eye, can be made visible, provided single phases are illuminated at very short intervals like lightning, with the length of the brief flashes of light being rather short in relationship to the duration of the whole period. When the frequency of the flashes exactly coincides with the frequency of the vibrations relative to the object, always the same vibratory phase is briefly illuminated, and the object, in reality rapidly vibrating, seems to stand still as seen by the eye ("standstill mode"). At which position the object seems to stand still is dependent on the fact at which vibratory phase the object is exposed to the light flashes. If the frequency of the flashes is rather close to the frequency of the vibration, but does not exactly coincide with it, in each passage not exactly the same position but the neighbouring one is illuminated. The vibratory cycle appears now to be slow, but this apparent cycle is composed of single images out of consecutive phases in consecutive periods forming to one continuous optical impression ("slow motion mode").

The stroboscopic principle - discovered in 1829 by the physicist Plateau from Brussels and by Stampfer from Vienna (1832/33), was used as early as in 1852, i.e. even before introducing the laryngeal mirror, by Harless for the first stroboscopic examinations with the excised larynx (7). The required pulse light was, at that time, provided with the aid of rotating perforated disks interrupting periodically or releasing a steady beam of light, and which were placed between the light source and the object to be examined. In the meantime, high-pressure photoflash lamps filled with noble gas and the flashing sequence of which is controlled electronically, have taken the place of the rotary disks after which the method is named (στροβος = turbulence). The frequency of the light flashes is triggered by the fundamental frequency of the voice which corresponds to the frequency of the vibrating vocal folds. In addition, high speed shutter cameras are being used to create the stroboscopic effect.

1.1.2 Laryngostroboscopy

Stroboscopy has been in use of laryngology for more than 125 years. In 1878, Oertel, for the first time ever, observed in Munich the larynx of a human being with the aid of a light source whose light flux was periodically released and interrupted through a rotating perforated disk and then, he saw the vocal folds vibrating. The first strobophotographic pictures Mueehold succeeded with in
Berlin, 1898 (the light sensitive plate of the camera had to be exposed several times because of the low intensity of the light source) are still admired wholeheartedly by everyone who ever has tried to freeze such images by photo. The breakthrough to clinical application was achieved in 1960 by Schoenhaerl with his monograph on Laryngostroboscopy (7).

1.1.3 Videolaryngo-Stroboscopy
The combination with both endoscopic and video technology is present day standard for the clinical application of stroboscopy in functional diagnostics of the larynx. Besides of new research technologies coming up for the investigation of the vibrating glottis, such as high speed glottography or videokymography, the application of the stroboscopic principle for functional diagnostics of the larynx is, for good reason, the basic and mostly used method to watch the glottal vibrations in an office setting of to-day (1, 2, 6, 14, 15). The optic illusion that provides us a slow motion like or standstill like picture of the acting vocal folds has proven to be a very useful tool. Compared to other techniques, as mirror examination, rigid or flexible endoscopy, stroboscopy was assessed highly superior regarding essential diagnostic information, and in combination with video recordings, no other method reached such high numbers of cases which required change of diagnosis (17), cf figure below. Thus, videolaryngostroboscopy can be recommended as a standard for clinical use without any reservation now as before.

![Fig. 1 Appreciation of methods: Laryngeal investigations (after Wuyts et al. 1995)](image)

After this brief introduction to the optic principles and the instrumental prerequisites as suitable for the ENT practitioner, the first section of the Instructional Course will, primarily, concentrate on conditions and procedures of examination, valuation and documentation of stroboscopic findings, sensitivity and specificity of the method, and indications for stroboscopic investigations. Video clips will be used for demonstration and refined evaluation of typical findings, combined with some basic acoustic analyses.
1.2 Basics of Glottal Biomechanics

1.2.1 Vocal Folds Vibration and Voice Sound
The relations between vocal fold vibration and the voice sound can be schematized, in a first approach, as follows:

<table>
<thead>
<tr>
<th>Vocal Folds Vibration</th>
<th>Voice Sound Perception</th>
</tr>
</thead>
<tbody>
<tr>
<td>frequency</td>
<td>pitch</td>
</tr>
<tr>
<td>amplitude</td>
<td>loudness</td>
</tr>
<tr>
<td>phases</td>
<td>sound colour, timbre</td>
</tr>
</tbody>
</table>

1.2.1.1 Frequency
The frequency of the vocal fold vibrations corresponds to the basic pitch (fundamental frequency) of the complex voice sound. In males it may, all in all, vary from ca 65 to 500 Hz, or in pitches from C to c² (C2 to C5, American notation), in females from ca 130 to 1,000 Hz or pitches from c to c³ (C3 to C6). Within the range of the mean pitch of the speaking voice (habitual pitch) in males, the vocal folds vibrate some 100 to 130 times per second (habitual pitch from G to c) and in females 200 to 260 times (habitual pitch g to c³).

1.2.1.2 Amplitude
The amplitudes of the vocal fold vibrations increase in width with increasing loudness. Though this increase is relatively low in comparison to the growing subglottal pressure, the increase can be observed distinctly in most of the cases.

1.2.1.3 Mucosal wave
In addition to the basic movement of the vocal folds, a so-called mucosal wave is appearing owing to the anatomically caused (Reinke’s space) shifting of the covering mucosa towards the body of the muscle. While the muscle is still vibrating in the upper portion of the vocal fold edge in the laterally directed opening phase, a narrow mucosal edge occurs at the lower vocal fold edge, initiating the closing phase. In this way a double contour as an upper and lower lip of the vocal fold edge is formed (vertical phase shifting). The pliability of the mucosa and the covering mucus are important for the synchronization of the two vibrating folds. The extent of the mucosal waves follows the extent of the amplitudes.

1.2.1.4 Phases
Out of the individual phases of the vibratory cycle (opening phase, closing phase, open phase, closed phase), the closed phase appears particularly meaningful. The longer this phase is in duration within a vibratory cycle, the more acoustic power will be transformed from aerodynamic energy and the more partial tones will characterize the spectrum of the voice.

1.2.1.5 Interrelations
However, it must be pointed out that the acting mechanisms in realizing different pitches, loudness and timbres are closely linked to each other. As a matter of fact, it is practically not possible to isolate one from another except when considering most fundamental mechanisms.

So, for instance, the amplitudes as well as the mucosal waves are reduced with a rising pitch owing to the higher tension of the vocal folds (lengthening of the vocal folds by tension according to the external frame positioning of the larynx, increase in the anterior-posterior distance by tilting back the cricoid plate due to the contraction of the Mm. cricothyroidei, approximation of the
cricoid arch closer to the frontal part of the thyroid cartilage). As already mentioned, the duration of the closed phase is associated with the power of the voice, i.e. the open quotient (time ratio from the open to the closed glottis) decreases with increasing loudness. Loud phonation is not only marked by a higher sound pressure level on the whole, but also by a greater number of partial tones, i.e. the timbre changes towards a “stronger” voice. Finally, the phases also change as a function of pitch: With rising pitch the closed phases become shorter.

These characteristics such as frequency, amplitude, mucosal wave and phase can be, surely, observed as individual characteristics of the vocal fold vibrations, however, they do not influence the voice sound in an isolated manner, but in a way that can be considered complex only. Interpretation of the observable vibratory variables of the vocal folds is, additionally, complicated due to the fact that

- there is a wide variability of all these parameters in normal vibrations
- there are no simple relations between the physical measures of the voice sound and the corresponding perceptual impressions (for instance pitch related alterations of the perception of loudness, emergence of subjective tones etc.).

1.3 Basics of Laryngostroboscopic Investigations

1.3.1 Conditions and Procedures

Topical **anaesthesia** of the oropharynx if necessary to avoid gagging reflexes which may influence the muscle tonus in the larynx, too, and thus falsify the original vibration pattern.

- **pitches** (basic orientation):
  - c (130 Hz) for males and one octave above
  - c¹ (260 Hz) for females and one octave above
- **loudness**: low, medium, high
- **mode sequence**: slow motion, standstill.

1.3.2 Valuation and Documentation

Always with reference to pitch and loudness of phonation (to be noted generally), because of their essential influence on the vibration pattern of the vocal folds.

- **Amplitudes**: medium, small (short), large (wide); hints at equality of the two sides or side differences
- **Phonatory Standstill**: left/right, occasional/always, partial/total
- **Mucosal Waves**: medium, small (reduced), large (enlarged), none (missing)
- **Phase Differences**: none, 90°, 180° (parallel vibrations)
- **Closure**: middle/posterior part; medium, short (shortened), long (prolongated), none (missing)
- **(Irregularities)**: none, rare, frequent, permanent  *(no reliable assessment possible according to the trigger principle)*

1.3.3 Sensitivity and Specificity

- Highly sensitive but barely specific in organic dysphonias
- little sensitive and not specific in functional dysphonias

Any valuation always and only with reference to the overall clinical picture. It has to be stressed that, according to recent studies, verification as well as subclassification of functional dysphonias by means of stroboscopic investigations only have been overestimated in the past because of the wide variability of normal vibrations (6).
1.3.4 Indications

- **Functional dysphonias**
  In spite of the reservations concerning functional dysphonias, stroboscopy is indispensable in all cases of dysphonia to exclude even minimal organic lesions reliably.

- **Organic dysphonias**
  are the real domain of stroboscopy, particularly for early selection of malignant pathologies by detection of nonvibrating portions, *phonatory standstill*, of the vocal folds. This phenomenon is essentially important in the frame of systematic check-ups of patients suffering from chronic laryngitis. Video recordings provide best conditions for long term follow-up comparisons.

- **Phonosurgery**
  In all cases of phonosurgical interventions careful pre- and posttherapeutic stroboscopic documentation are a must.
  Furthermore, phonosurgical procedures themselves can gain accuracy considerably in many cases, when the interventions are carried out under indirect microstroboscopy, with the patient awake, under topical anaesthesia. This approach provides optimal functional control by visual and auditory monitoring while surgery is being done. Personal experiences from some 7,000 operations have proven the advantages of indirect laryngeal microsurgery with the aid of stroboscopic light.

- **Evidence based therapy**
  Stroboscopy is indispensable for efficacy evaluations of any kind of therapeutic measures applied to voice patients and the same is true for also in connection with

- **Medical expert reports**

1.4 Mediastroboscopy

Mediastroboscopy stands for an advanced digital media technology and provides a comfortable digital video system that is easy to handle and adapted to the everyday needs of the specialized practitioner. The main advantage of this system is a very transparent documentation and easy and quick access to any relevant detail of an even large data base. Additionally, basic acoustic analyses of the voice sound of recorded stroboscopic video clips are available. All data may be exported for insertion into other documents, as there are records or reports. The system can be used for both general endoscopic examinations as well as for situations in which a stroboscope is used. The course will focus on the stroboscopic approach and demonstrate the main features of the system.

1.4.1 New Patients

For new patients, a new session will be created and the recorder mode is started. Displays for fundamental frequency (Hz) / musical pitch and sound pressure level (dB(A)) / 5 subjective loudness classes (<1, 1, 2, 3, >3) allow precise control of the phonatory conditions under the investigation for appropriate clinical interpretation of the vibratory pattern. After recordings, the play mode offers endless loops and freeze function to scrutinize the selected clips concerning pertinent findings. Relevant video clips may be selected and entered into the data base by just pressing the button [Accept]. With the editing mode, personal data as well as findings' descriptions, diagnoses, or comments can be added, in the course of which predefined text modules are available for common findings and diagnoses to be inserted. Individual pictures may be printed or exported for insertion into other documents.
1.4.2 Archive Data
Lists for patient names or numbers present the archive data from which can be selected by search buttons (search ABC/123). Then the procedure continues likewise as described above for new patients.

1.4.3 Basic Acoustic Analyses
The attached program “ATMOS Voice 03” provides the time signal, fundamental frequency distribution, jitter, shimmer, LPC, sonagram, FFT and some other data which can all be displayed for each of the sound tracks of the recorded video clips for further evaluation, in order to put voice data and vibration data of the vocal folds together. In addition, real time phonetography is available. This is the topic for the second section of the Instructional Course.

2 Phonetography

2.1 Introduction
2.1.1 Ranges of Pitch and Loudness
In addition to findings by stroboscopy or other investigations, voice range profiles represent a clear and reproducible description of basic properties of the voice (3, 8, 9, 10, 12).

The phonetogram as the result of a two dimensional measurement of the range of the fundamental frequency (F₀) for pitch, displayed against the range of the sound pressure level (SPL) for loudness, shows the dynamic capacities of a voice at a glance. In a coordinate system, musical pitch or fundamental frequency (Hz) is displayed on the abscissa, SPL (dB(A)) for loudness on the ordinate.

2.1.2 Singing and Speaking Voice Profiles
With the phonetogram, several capacities of a voice can be examined separately. The singing voice profile represents the vocal function for sustained tones. Each tone of a voice range should be sung with minimum and with maximum loudness possible. In a normal voice profile, an ellipsoid figure with ascending longitudinal axis appears on the diagram. Depending on gender, voice type, existing voice disorders and the reliability of the measurements, a great variety of figures may be possible.

The relation between fundamental frequency (Hz) and SPL (dB(A)) for running speech is recorded in the speaking voice profile. Speaking (e.g. counting) in soft, colloquial, lecturing, and (as loud as possible) shouting voice mode is measured and indicated on the diagram by crosses.

Critical data of interest are the average pitch in normal speaking (mean speaking frequency) and the maximum SPL reached in shouting.

For the singing voice profile, an average of two octaves for the pitch range and 50 dB(A) for the SPL range can be expected in cases of normal, healthy voices. This corresponds to a maximum SPL of, at least, 90 dB(A) in the speaking voice profile. As to the mainly used voice features, they cover, for females, the pitch range g - e² and an SPL range of 25 - 30 dB(A), and for males A - e¹ and 25 – 35 dB(A) respectively (11). The normal mean speaking frequency (habitual pitch) is expected at the transition from the lower to the middle third of the total pitch range or, in terms of musical intervals, a fourth to a fifth above the lowest tone of the pitch range, i.e. G – c for males and g – c¹ for females (in Europe, a tendency can be observed that this pitch is shifting a minor third lower to e – a).
2.1.3 Dysphonia Severity Index (DSI)

To date, objective valuation of dysphonia is, still, not resolved adequately. Too many parameters are influencing the properties and the efficiency of voice. The subjective, perceptual assessment of hoarseness in terms of a 4 point scoring of severity (0 = none/normal, 1 = mild, 2 = moderate, 3 = severe) as used for the GRBAS (1) as well as for the RBH scale (4, 16) was chosen as reference to an objective measure, attempting to reflect main dimensions of voice properties by one numeral, the dysphonia severity index (DSI). It is based on a weighted combination of selected measurable parameters, including phonetographic data, with important impact on overall vocal capacity (18):

- Maximum phonation time (MPT) in s
- Highest fundamental frequency (F₀-High) in Hz
- Lowest intensity (I) in dB(A) for SPL
- Jitter in %

\[ \text{DSI: } 0.13 \times \text{MPT (s)} + 0.0053 \times F₀-\text{High (Hz)} - 0.26 \times I-\text{low (dB(A))} - 1.18 \times \text{Jitter (dB)} + 12.4 \]

The index was derived from multivariate analysis of the data measured and stepwise analysis methods, and finally constructed as to the results of a so-called proportional odds logistic regression.

The average DSI reaches from +5 for normal voices to -5 for severely hoarse voices. The worse the patient’s voice quality, the more negative becomes the DSI. In regard to the subjective, perceptual 4-point grading of hoarse voice qualities according to the RBH scale, a reliable relation to the corresponding objective DSI values has been verified (5).

![Fig. 2 Means and standard error of means of the DSI versus hoarseness degree (Nawka)](image-url)
2.2 Basics of Phonetographic Investigations
2.2.1 Conditions and Procedures (Standards)
According to the recommendation by the Union of the European Phoniatricians (UEP), the following standards should be taken into consideration (8):
A simple sound level meter and a tone generator are required, A-weighted network (dB(A)) is advised, combined with slow meter damping as minimum instrumental equipment.
The microphone should be placed in a distance of 30 cm to the speaker and ought to have omnidirectional characteristics. The commonly present surrounding noise may not exceed 40 dB(A). Measuring should be done in rooms with ‘normal room acoustics’ or sound treated rooms with moderately damped reverberation.
To be accepted for registration, a minimum of 2 s per phonation is set as basic time standard. Registration of the measuring results on a standardized form is strongly recommended. For the paper size DIN A5, 15 mm for 10 dB (range: 40-120 dB (A)) on the ordinate, and 36 mm for one-octave (range: G1 – g3) distance on the abscissa were defined.

2.2.2 Valuation and Documentation
The experienced user will draw basic information at a glance by just looking at the graph. In addition, he may read out essential data in terms of numerical figures, as there are highest frequency, lowest frequency, frequency range, maximum SPL, minimum SPL and SPL range. For documentation, prints can be added to the patients’ records and the data can be stored on a computer.

2.2.3 Sensitivity and Specificity
Phonetographic data are highly sensitive concerning vocal capacities and may be interpreted as basic manifestation of vocal efficiency. But, this does not imply any sensitivity with regard to nosological differentiations. Different voice disorders may show all just quite the same phonetograms, and from the same type of voice disorders quite different phonetograms may be recorded. Thus, no specificity can be expected in relation to specific diseases either.

2.2.4 Indications
With the above mentioned reservations in mind, assessment and documentation of basic voice parameters form the main indication. These important aspects reflecting vocal efficiency are required for follow-up purposes and in all cases of expert opinions covering
- All voice disorders
- Pre-posttherapeutic comparisons
- Voice aptitude tests
- Singers’ voice classification

2.3 The LingCom lingWAVES Phonetogram Pro
2.3.1 Technical Prerequisites
Personal computer from Pentium IV, Windows 2000, XP.
Soundblaster compatible soundcard, USB (for Dongle), USB/RS232 data linked sound level meter (included in the package)
Hard disk needs: installation interface approx. 30 MB, working storage: minimum 512 MB, depending on the size of the sound files additional working storage is needed.

2.3.2 Calibration
The lingWAVES Phonetogram Pro uses a data linked sound level meter for a high quality digital measurement of the sound pressure level. There is no need for a sound pressure calibration. The
sound level meter microphone is also used to get speech signal data for fundamental frequency calculation.

2.3.3 Special Features

2.3.3.1 lingWAVES Phonetogram

With the phonetogram module of lingWAVES Phonetogram, real time measurement of fundamental frequency and sound pressure level is possible. The voice range, recorded by microphone and soundcard, is illustrated automatically in a two dimensional polygon diagram (phonetogram). A coloured cluster view mode offers a special display for measurement of the singing voice.

2.3.3.2 lingWAVES Phonetogram Pro

Additional features are available in the lingWAVES Phonetogram Pro module. An easy and vivid illustration of speaking voice in different modes (soft, normal, loud) should be a frequently used measurement. It is possible to toggle between the cluster and the polygon view of the diagrams to compare different phnetograms, or to combine measurements of singing and speaking voice. Especially the examination of professional singers often includes assessment of voice quality. An important aspect of a penetrating voice is the share of high frequency formants (singers formant, representing the “ring” of a voice) of the total sound of voice. While recording the singing voice with the Phonetogram Pro module, it is possible, to show the SPL of the high frequency formants in the area of 2,500 to 5,000 Hz simultaneously.

2.3.3.3 Dysphonia Severity Index (DSI)

Automatic setting of DSI is possible with the Phonetogram Pro module. In addition to required measurements of highest frequency (F₀-High in Hz) and lowest intensity (I-Low in dB), which are registered while recording the phonetogram, maximum phonation time (MPT in s), and Jitter (%) can be added easily. The DSI will be calculated by the program and displayed on the graph.

Fig. 3 Polygon view of the voice range profile. Dotted line: High formant.

Crosses indicating speaking voice measures for soft, normal and loud speaking

The Phonetogram Pro module also comprises a simple voice strain test (loading test), that displays the pitch and SPL of the voice in running speech (reading) over a certain period of time. The properties of this test regarding target SPL and duration can be set according to standards or individual requirements. Changes in pitch and/or SPL may be symptoms of pathologic voice tiredness.
2.3.1.4 lingWAVES modular concept

The lingWAVES software is based on a modular concept (Plug-In). It is easy to add additional analysis and processing modules (e.g. voice signal package, The Göttingen Hoarseness Diagram) while using the same user dialog software.

The course, sponsored by ATMOS, is focused on office requirements of everyday practice and will be illustrated by video demonstrations from DVD.

References

7. Schoenhaerl E: Die Stroboskopie in der praktischen Laryngologie. Thieme, Stuttgart 1960
15. Wendler J: Stroboscopy – Principles and clinical application in the investigation of the larynx. ATMOS Medizintechnik, Lenzkirch 1993

Addresses of the authors

Prof. Dr. med. Jürgen Wendler, An der Wuhlheide 230 E, D-12459 Berlin, Germany
Email: juergen.wendler@charite.de

Prof. Dr. med. Tadeus Nawka, Abt. für Phoniatrie und Pädaudiologie, HNO-Klinik, Ernst-Moritz-Arndt Universität Greifswald, Wallther-Rathenau-Str. 43-45, D-17487 Greifswald, Germany
Email: nawka@uni-greifswald.de

Dr. med. Dirk Verges, Abt. für Phoniatrie und Pädaudiologie, HNO-Klinik, Charité – Universitätsmedizin (Campus Mitte), Schumannstr. 20/21, 10117 Berlin, Germany
Email: dirk.verges@charite.de